BMJ Open

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research programme: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease at an Australian tertiary paediatric hospital

Michael J Coffey ![ORCID],[1] Isabelle R McKay,[1] Michael Doumit,[2] Sandra Chuang,[1,3] Susan Adams,[1,4] Sacha Stelzer-Braid,[5,6] Shafagh A Waters,[1,7] Nadine A Kasparian,[1,8] Torsten Thomas,[9] Adam Jaffe,[1,10] Tamarah Katz,[11] Chee Y Ooi[1,10]

For numbered affiliations see end of article.

Correspondence to
Dr Chee Y Ooi;
keith.ooi@unsw.edu.au

ABSTRACT

Introduction Chronic gastrointestinal and respiratory conditions of childhood can have long-lasting physical, psychosocial and economic effects on children and their families. Alterations in diet and intestinal and respiratory microbiomes may have important implications for physical and psychosocial health. Diet influences the intestinal microbiome and should be considered when exploring disease-specific alterations. The concepts of gut-brain and gut-lung axes provide novel perspectives for examining chronic childhood disease(s). We established the 'Evaluating the Alimentary and Respiratory Tracts in Health and disease' (EARTH) research programme to provide a structured, holistic evaluation of children with chronic gastrointestinal and/or respiratory conditions.

Methods and analysis The EARTH programme provides a framework for a series of prospective, longitudinal, controlled, observational studies (comprised of individual substudies), conducted at an Australian tertiary paediatric hospital (the methodology is applicable to other settings). Children with a chronic gastrointestinal and/or respiratory condition will be compared with age and gender matched healthy controls (HC) across a 12-month period. The following will be collected at baseline, 6 and 12 months: (i) stool, (ii) oropharyngeal swab/sputum, (iii) semi-quantitative food frequency questionnaire, (iv) details of disease symptomatology, (v) health-related quality of life and (vi) psychosocial factors. Data on the intestinal and respiratory microbiomes and diet will be compared between children with a condition and HC. Correlations between dietary intake (energy, macro-nutrients and micro-nutrients), intestinal and respiratory microbiomes within each group will be explored. Data on disease symptomatology, quality of life and psychosocial factors will be compared between condition and HC cohorts.

Strengths and limitations of this study

► The prospective, longitudinal, controlled, observational design of this research programme provides a structured approach which can be simultaneously applied to multiple chronic gastrointestinal and/or respiratory conditions of childhood and utilises a universal control cohort (for age and gender matching).

► This study will simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes, which will tease out disease-causing alterations in the microbiomes, provide insights into the gut-lung axis and potentially identify modifiable dietary factors.

► We will explore relationships between the primary outcomes (diet, intestinal and respiratory microbiomes) and health-related quality of life (including symptomatology), which may provide insights into the gut-brain axis and identify novel pathogenic mechanisms in these conditions.

► A limitation of this research programme is that it currently includes a single centre, Sydney Children's Hospital Randwick, Australia, however it is a tertiary referral centre for a diverse group of children across the state of New South Wales, Australia.

► A further limitation is the arbitrary sample size targets given the exploratory nature of these studies.

Results will be hypothesis-generating and direct future focussed studies. There is future potential for direct translation into clinical care, as diet is a highly modifiable factor.

Ethics and dissemination Ethics approval: Sydney Children's Hospitals Network Human Research Ethics

INTRODUCTION

The primary disease burden in childhood has shifted over the last century from infectious to chronic diseases.[1] Chronic childhood diseases, encompassing a wide spectrum of conditions with different pathogeneses, may have long-lasting physical, psychosocial and economic effects on children and their families.[2][3] The human microbiome is a collection of all microorganisms (bacteria, viruses, archaea and eucaryotes) living in association with the human body.[4] Our understanding of the human microbiomes in health and disease has begun to develop due to the advent of high-throughput sequencing and mass-spectrometry technologies, with the gut emerging as an ecosystem of particular interest. While the effects of an altered gut microbiome (dysbiosis) may not apply to all chronic diseases, there are conditions, disease-related complications and comorbidities linked to gut microbial dysbiosis. This is especially true in chronic gastrointestinal and respiratory conditions. Affected children are at risk of an imbalanced diet as well as mental health difficulties, which in turn can influence eating behaviours, attitudes and nutritional intake.[5][6] Some of these conditions also require lifelong dietary modifications; for example, cystic fibrosis (CF).[7] Additionally, the complex interaction between microbiota (ie, bacteria), available nutrients and the immune system is essential in maintaining homoeostasis and fighting against invading pathogens at mucosal sites.[8] An important limitation common to most current publications on the human intestinal microbiome in chronic childhood disease(s) is the lack of quantifiable dietary data, as the diet has a marked influence on gut microbiota in health.[9]

The principles and framework of this research programme were developed to be applicable to many chronic gastrointestinal and/or respiratory conditions of childhood. Due to the clinical and/or research expertise of the authors and for the purposes of this manuscript, we will describe this programme based on three relevant chronic diseases: (i) CF, (ii) obstructive sleep apnoea (OSA) and Hirschsprung's disease (HSCR). These conditions all have reported or expected changes in their intestinal and/or respiratory microbiomes.

CF is the most common life-shortening recessive disease in Caucasians.[10] It is characterised by intestinal malabsorption, impaired growth and nutrition and lung disease.[11] In CF, a high calorie, high fat diet (110% to 200% of recommended daily energy intake) is advised to prevent malnutrition and optimise growth.[7] Recent reports suggest that children tend to achieve the recommended CF diet primarily by overconsumption of energy-dense, nutrient-poor foods rather than nutrient-dense foods.[12] We have previously reported that children with CF, from as early as infancy, have alterations in their gut microbiota, impaired innate immunity and intestinal inflammation.[13–17] We have also observed that poor growth in children with CF is significantly correlated with the degree of intestinal inflammation.[14] The aetiology of gut microbial dysbiosis and inflammation in CF remains unclear. It is plausible that dietary intake plays a role, as enteric fat abundance (from a high-fat diet) may select for a pro-inflammatory microbiota.[18][19] Alterations in intestinal metabolomic[20] and proteomic[21] profiles have also been reported. As the life expectancy of CF patients improves, age-related diseases such as gastrointestinal malignancies and cardiovascular disease (eg, myocardial infarcts in adults with CF) are a growing concern.[22] Thus, optimal strategies to optimise health and reduce disease risk factors need to be determined.

In children, OSA can have cardiovascular, neurocognitive and behavioural consequences.[23] Murine studies suggest intermittent hypoxia, hypercapnia and sleep fragmentation promote intestinal dysbiosis, increased visceral fat mass, systemic inflammation and atherosclerosis.[24–27] Additionally, the inhibition of gut microbial metabolites attenuating atherosclerosis[26] and replication of hypertension after faecal transplant from hypertensive to normotensive OSA rats[28] suggest the possibility of influencing clinical outcomes through affecting the gut microbiome. In adult studies, OSA is associated with gut epithelial damage,[29] and nasal dysbiosis and inflammation.[30]

HSCR is a congenital disorder where the distal intestine is aganglionic for a variable length. This results in a functional bowel obstruction that usually presents in newborns. Following corrective surgery, children often have ongoing intestinal symptoms, and Hirschsprung-associated enterocolitis (HAEC) remains the most frequent complication. This may result in frequent hospitalisations and even mortality. Children with and without HAEC often have an altered intestinal microbiome[31] and altered composition of short chain fatty acids.[32]

To the best of our knowledge, there are no publications on the intestinal virome (ie, viruses) in children with CF, OSA or HSCR. Bacteriophages (viruses which infect bacteria) can influence bacterial populations via host lysis and horizontal gene transfer, as well as indirectly regulate immune function and inflammation.[33–36]

Despite accumulating evidence linking health, diet and the microbiomes, there is a paucity of research exploring this simultaneously in the context of chronic paediatric disease.[37] Furthermore, potential gut-brain[38] and gut-lung[39] axes have yet to be well characterised in these conditions. The gut-brain axis refers to the bidirectional communication between the central nervous systems and gut microbiome, and is mediated by neural, endocrine and immune pathways.[40] The gut-lung axis refers to the bidirectional relationship between the gut and lungs, as there appears to be an immunological relationship between them.[39][41] Simultaneous, longitudinal studies using an integrated 'omics' approach will help to identify the functional consequences and pathogenic

mechanisms that occur within the altered intestinal and respiratory milieu in chronic conditions. By exploring disease mechanisms and environmental interactions (eg, diet) we may in turn develop insights into potential therapeutic strategies. Additionally, we may be able to identify whether diet may be amenable to specific modifications which may in turn benefit the intestinal microbiome.

The EARTH programme has been established to provide a structured approach to analysing the gastrointestinal and respiratory microbiomes and diet in children with a chronic gastrointestinal and/or respiratory condition. The design improves efficiency by recruiting and assessing a healthy control (HC) group which can be used for comparison against each of the conditions (as opposed to recruiting a new HC group for each condition). Although our initial design is focussed on CF, OSA and HSCR, the programme framework is applicable to other chronic gastrointestinal and/or respiratory conditions of childhood.

Objectives

The objective of this research programme is to evaluate and compare children with a chronic gastrointestinal and/or respiratory condition and age and gender matched HC. The primary objectives include analysing the intestinal and respiratory microbiomes (using an integrated 'omics' approach) and dietary intake using validated, parent-report tools (table 1). The secondary objectives are also presented in table 1 and include evaluating:

1. Known inflammatory biomarkers.
2. Symptomatology and health-related quality of life (HRQOL) using validated measures.
3. Phenotypic and clinical information.
4. Sociodemographic factors.

Additional secondary objectives include correlating within children with the same condition: (i) dietary intake with the intestinal microbiome, (ii) dietary intake with the respiratory microbiome and (iii) the intestinal and respiratory microbiomes.

We hypothesise that:

i. Children with chronic gastrointestinal and/or respiratory conditions will have altered intestinal and respiratory microbiomes compared with healthy children, and

ii. Diet plays a key role in influencing the intestinal and respiratory microbiomes and this may impact on clinical outcomes, biomarkers of disease and health-related quality of life.

To our knowledge, this programme will enable the first series of studies comparing the intestinal and respiratory microbiomes and diet in children with chronic gastrointestinal and/or respiratory conditions. Initial results will be hypothesis-generating and used to direct future studies tailored to a specific focus or line of inquiry. Additionally, studies from this research programme have potential for direct translation into clinical care as diet is a highly modifiable factor.

METHODS AND ANALYSES
Study design

The EARTH programme provides a framework for a series of prospective, longitudinal, controlled, observational studies, with each individual study comparing children with a chronic gastrointestinal and/or respiratory condition to HC. A single healthy control group will be used for comparison against all conditions and healthy controls are defined as children who are free of any chronic disease. The standardised methodological approach will also allow for comparisons between different health conditions. The Standard Protocol Items: Recommendations for Interventional Trials reporting guidelines were used for this protocol.[42]

Setting

Studies will be carried out at a single centre; the Sydney Children's Hospital (SCH) in Randwick, Australia. SCH is a tertiary paediatric hospital.

Participants

Children are eligible if they:
► Are aged between 0 and 18 years;
► Have been diagnosed with a chronic gastrointestinal and/or respiratory condition defined by consensus diagnostic criteria; or
► Are free of any chronic health condition (healthy control group); and
► Have a parent(s)/carer(s) who provides informed consent, or are at least 16 years old and provide informed consent.

Ineligibility criteria include:
► Children with more than one concurrent or unrelated chronic disease;
► Inability to comply with study requirements;
► Parent(s)/guardian(s) are unable to speak English or do not have a reading level age of at least 12 years.

Participants with a chronic gastrointestinal and/or respiratory condition will be matched to a HC for gender and age (as closely as possible).

Recruitment strategy

Participants with chronic gastrointestinal and/or respiratory conditions will be approached at their routine clinical appointments in the outpatient department. Flyers will be placed in the hospital for recruitment of HC. Prior to study participation, detailed written and verbal information will be provided about the content and extent of the study. Written informed consent from the parent/legal guardian of each participant will be required. If the child is deemed Gillick competent,[43] they will be encouraged to sign a specific child assent form. Parents/legal guardians and participants may withdraw consent at any time.

Outcome measures

The outcomes measures are presented in table 1. All samples, questionnaires and data will be collected from all participants at each time point. Presented below

Table 1 Primary and secondary objectives with related outcome measures

Domain	Data source	Technique	Outcome measures	Between group analyses*	Within group analyses†
Primary objectives					
(1) Intestinal microbiome	(1) Stool sample	Bacterial communities (16S rRNA (V4)[13] or MSS)	Alpha diversity (richness and Shannon index)	Student's t-test or Wilcoxon signed-rank test	Pearson or Spearman correlations with: ▲ Gastrointestinal microbiome ▲ Respiratory microbiome ▲ Diet ▲ Secondary objectives Descriptive
			Beta diversity (UniFrac) distances[46]	PERMANOVA[73]	
			Relative abundances of bacteria	ANCOM[74]	
		Viral communities (metagenomic sequencing[49])	Alpha diversity (richness and Shannon index)	Student's t-test or Wilcoxon signed-rank test	
			Beta diversity (Bray-Curtis dissimilarities)	PERMANOVA[73]	
			Relative abundances of viruses	ANCOM[74]	
(2) Respiratory microbiome	(2) Oropharyngeal swab or sputum sample	Proteomics (LC-MS)[21]	Protein z-score normalised LFQ intensities	Student's t-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
		Metabolomics (UHPLC-MS/MS)[20]	Metabolite normalised abundance	Student's t-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
(3) Diet	(i) ACAES (ages 2 to 18year)[54-56] (ii) 24 hours food recall (ages 0 up to 2year)		Energy intake	Student's t-test, Wilcoxon signed-rank test or Fisher's exact test	
			Percent energy from core foods		
			Macronutrient intake		
			Micronutrient intake		
			Diet quality score‡		
Secondary objectives					
(1) Biomarkers	Stool, oropharyngeal swab or sputum sample	ELISA	Inflammation (calprotectin, M2-PK, CRP and interleukins)	Student's t-test, Wilcoxon signed-rank test or Fisher's exact test	Descriptive
(2) Symptomatology and HRQOL	PedsQL Infant Scales (0–2 year) and gastrointestinal symptoms module (2–18year)[54-56]		HRQOL and gastrointestinal symptoms	Student's t-test, Wilcoxon signed-rank test or Fisher's exact test	Descriptive
	Rome IV Questionnaire[57 58]		Gastrointestinal symptoms		
	Spence Children's Anxiety Scale[59 60]		Anxiety symptoms		
	Short Mood and Feelings Questionnaires[61 62]		Depressive symptoms		
(3) Phenotypic and clinical information	Anthropometrics		Z-scores; weight, length/height, weight-for-length (ages 0 to 2year) and BMI (ages 2 to 20year)	Student's t-test, Wilcoxon signed-rank test or Fisher's exact test	Descriptive
	Clinical presentations		Number and length of hospitalisations, emergency department presentations, medications, vaccinations		
	Results		Biochemistry, microbiology and imaging results		
	Perinatal factors		Mode of delivery, feeding during infancy		
(4) Socio-demographic factors			Ethnicity	Descriptive	Descriptive
			SEIFA code[63]	Descriptive	Descriptive

All samples, questionnaires and data will be collected from all participants at each time point.
*Between group analyses describe comparisons between a condition and healthy control groups.
†Within group analyses describe analyses of two outcome measures within subjects of the same condition group.
‡ACAES only.
ACAES, Australian Child and Adolescent Eating Survey; ANCOM, analysis of composition of microbiomes; BMI, body mass index; CRP, C-reactive protein; HC, healthy control; HRQOL, health-related quality of life; LC-MS, liquid chromatography-mass spectrometry; LFQ, label-free quantification; M2-PK, M2 pyruvate kinase; MSS, metagenomic shotgun sequencing; PERMANOVA, permutational multivariate analysis of variance; rRNA, Ribosomal nucleic acid; SEIFA, socio-economic indexes for areas (a measure of relative socio-economic advantage and disadvantage in Australia); UHPLC-MS/MS, ultra-high performance liquid chromatography-tandem mass spectrometry; V4, Hypervariable region V4.

is a simplified explanation of each outcome/variable included in the research programme.

Primary outcomes/variables

1. Intestinal microbiome assessed from a stool sample using one or more of:
i. Bacterial community analysis (16S ribosomal RNA (V4)[13] or metagenomic shotgun sequencing):
   a. Alpha diversity indices:
      – Richness: the total number of unique species.[44]
      – Shannon index: a measure of both species abundance and evenness.[45]
   b. Beta diversity indices:
      – UniFrac: a distance metric used to compare biological communities that incorporates phylogenetic distances between observed organisms.[46]
      – Bray-Curtis dissimilarity: a count metric used to quantify the compositional dissimilarity between two different sites.[47]
   c. Relative abundance: the per cent composition of an organism relative to the total number of organisms in the area.[48]
ii. Viral community analysis (metagenomic sequencing),[49] as above for bacterial community analysis.
iii. Proteomics (liquid chromatography-mass spectrometry (LC-MS)):[21]
   a. Protein z-score normalised label-free quantification intensities.
   b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.
iv. Metabolomics (ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS)):[20]
   a. Metabolite normalised abundance.
   b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.
2. Respiratory microbiome assessed from an oropharyngeal swab or sputum sample, using one or more of the techniques listed above (1a–d).
i. A sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate.
   a. Associations with the intestinal microbiome (1) will be used to explore the gut-lung axis.
3. Dietary intake assessed using a validated semi-quantitative food frequency questionnaire:
i. For participates aged 0 up to 2 years, a 24 hours food recall:
   a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake) and micronutrient intake (total intake and proportion of energy intake).
ii. For participants aged 2 to 18 years, the Australian Child and Adolescent Eating Survey (ACAES).[50–53] The ACAES is a validated food frequency questionnaire (120 items, semi-quantitative) used to quantify food and nutrient intake over the preceding 6 months,

developed and validated for use in Australian children:
   a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake), micronutrient intake (total intake and proportion of energy intake) and overall diet quality score.

Secondary outcomes/variables

1. Faecal and respiratory inflammatory biomarkers, such as calprotectin, M2 pyruvate kinase, C-reactive protein and interleukins.
2. Symptomatology and health-related quality of life (HRQOL) will be collected directly from children where age-appropriate measures exist and/or parents using age-appropriate measures:
   a. PedsQL Infant Scales (0 to 2 year) and gastrointestinal symptoms module (2 to 18 year)[54–56] (HRQOL and gastrointestinal symptoms).
   b. Rome IV Questionnaire[57 58] (gastrointestinal symptoms). Designed to diagnose functional gastrointestinal disorders, which are defined as disorders of the gut-brain interaction in children aged 0 to 18 years. These criteria capture gastrointestinal symptoms which are relevant to motility disturbance, visceral hypersensitivity, altered mucosal and immune function, altered gut bacteria and altered central nervous system processing;
      i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
   c. Spence Children's Anxiety Scale[59 60] (anxiety symptoms);
      i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
   d. Short Mood and Feelings Questionnaires[61 62] (depressive symptoms).
      i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
3. Anthropometrics, including z-scores for weight, length/height, weight-for-length (ages 0 to 2 years) and body mass index (BMI) (ages 2 to 20 years).
4. Z-scores; weight, length/height, weight-for-length (ages 0 to 2 year) and BMI (ages 2 to 20 year).
5. Clinical information and biochemical results obtained through routine care, such as number and length of hospitalisations, emergency department presentations, perinatal factors (mode of delivery, feeding type(s) in infancy), medications, vaccination status (including timing of most recent vaccination), biochemistry, microbiology and imaging results.
6. Sociodemographic factors such as ethnicity and socioeconomic indexes for areas code.[63]

Procedures

Each participant will be assessed on three occasions over a 12-month period; at study entry, 6 and 12 month follow-up. At each time point, the following will be collected:
► A stool sample;

- ► An oropharyngeal swab or sputum sample (a sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate);
- ► Dietary intake measured using the ACAES (2 to 18 years)or 24 hours food recall (0 up to 2 years);
- ► A secure, password-protected online survey comprising:
  i. PedsQL Infant Scales (0 to 2 year) and gastrointestinal symptoms module (2 to 18 year),[54–56] tailored to age;
  ii. Rome IV Questionnaire[57 58] (0 to 18 years);
  iii. Spence Children's Anxiety Scale[59 60] (3 to 18 years);
  iv. Short Mood and Feelings Questionnaires[61 62] (6 to 18 years);
  v. Clinical and biochemical results obtained through routine care and hospitalisations (if available);
  vi. Sociodemographic factors (baseline survey only);
- ► Anthropometrics: height, weight and BMI z-scores.

Details regarding sample and data collection

Stool samples will be collected in a sterile specimen jar using a Feces Catcher (Abbexa Ltd, Cambridge, UK). Sputum samples will be collected from children who are able to expectorate and oropharyngeal swabs will be collected in those children who cannot expectorate sputum. All samples will be transported in a cooler bag (with a −18°C ice pack) to the hospital laboratory within 24 hours of collection. Stool, oropharyngeal swab and sputum samples will be aliquoted and stored at −80°C. Where practical, samples will be collected at least 4 weeks after completion of oral/intravenous antibiotic therapy (excluding prophylactic antibiotics) or an acute infectious illness.

For children aged 0 up to 2 years, a 24 hours dietary recall will be conducted by a study dietician. For children aged 2 to 18 years, the ACAES will be completed online by parents/guardians or by the child themselves if aged over 14 years.

Participant clinical information and questionnaires will be collected using the Qualtrics online survey software tool (www.qualtrics.com), which is distributed via a secure email link. Qualtrics is a secure, password-protected platform, which allows for the distribution of electronic surveys and collection of data. The surveys are programmed to be condition/control-specific and facilitate a personalised flow depending on the time point of the study and the age of the participant (to facilitate administration of age-appropriate questionnaires). Further details on each of the measures used for evaluating symptomatology and HRQOL are presented in table 2. Clinical and biochemical results will be obtained via the SCH electronic medical record system and recorded in a Qualtrics survey.

Participants will be measured for their height and weight using standardised methods.

Sample and data processing techniques

Processing of stool, oropharyngeal swab and sputum samples is almost identical (sparing a few initial sample preparation steps). For bacterial community analysis, DNA will be extracted using QIAamp DNA kits (QIAGEN, Hilden, Germany) according to manufacturer's instructions. For 16S rRNA gene analysis specifically, amplification will be performed with primers 515F and 806R spanning the V4 region and sequencing data will be processed using USEARCH.[64]

In the instance where species resolution of bacterial communities is thought to be beneficial, metagenomic shotgun sequencing (MSS) will be performed as an alternative to 16S rRNA gene sequencing. For MSS, no amplification step will be performed prior to sequencing. Sequencing data will be processed using a custom in-house pipeline.

For viral community analysis specifically, sample preparation will follow an adjusted NetoVIR (Novel Enrichment Technique Of VIRomes) protocol.[49] All sequencing will be performed using the Illumina MiSeq platform at the Ramaciotti Centre for Genomics at the University of New South Wales (UNSW). Briefly, sequencing data will be processed using the Vipie platform[65] for taxonomic assignment and VirSorter pipeline[66] for functional annotation.

For untargeted proteomics, samples will undergo an adjusted Debyser *et al* protocol for protein extraction, gel electrophoresis and analysed using LC-MS/MS at the Bioanalytical Mass Spectrometry Facility (BMSF), UNSW.[21] Briefly, proteomics data will be analysed using MaxQuant[67] and Ingenuity Pathway Analysis (QIAGEN).

For untargeted metabolomics, metabolites will be extracted in 1:1 (v:v) acetonitrile:H_2O and analysed using a U3000 UHPLC system coupled to a Q-Exactive mass spectrometer (MS; Thermo Fisher Scientific) at the BMSF, UNSW. Briefly, metabolomics data will be analysed using Progenesis CoMet (Waters/NonLinear Dynamics).

Faecal and respiratory biomarkers (listed above) will be measured using ELISA.

Nutrient intake data from the ACAES and 24 hours recall is computed using FoodWorks (V.3.02.581) and the following databases: Australian AusNut 1999 database (All Foods) Revision 14 and AusFoods (Brands) Revision 5 (Xyris Software (Australia) Pty Ltd, FoodWorks Professional V.3.02.581. 2004: Brisbane Australia). Outputs include a quantified estimate and the percentage of energy from a wide of range of macro-nutrients (protein, fat, carbohydrate) and micro-nutrients (vitamins A, B, C and minerals such as iron, zinc and calcium). In addition, overall diet quality score and the percentage of energy derived from nutrient rich core foods and energy-dense, nutrient-poor discretionary foods is calculated.

Administration of patient records and data

At the time of consent and enrolment, participants will be assigned a unique study ID number (nine alphanumeric characters). All patient records, samples and data

Table 2 Measures for symptomatology and health-related quality of life (HRQOL)

Measure	Domains (Items)	Scoring	Interpretation
(i) PedsQL			
Infant Scales – parent report for infants (ages 1–12 months)[56]	Total (36): 5-point LS. Physical functioning (6), physical symptoms (10), emotional functioning (12), social functioning (4), cognitive functioning (4).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
Infant Scales – parent report for infants (ages 13–24 months)[56]	Total (45): 5-point LS. Physical functioning (9), physical symptoms (10), emotional functioning (12), social functioning (5), cognitive functioning (9).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
3.0 Gastrointestinal symptoms module – parent report for toddlers (ages 2–4)[54 55]	Total (74): 5-point LS. Stomach pain and hurt (6), stomach discomfort when eating (5), food and drink limits (6), trouble swallowing (3), heartburn and reflux (4), nausea and vomiting (4), gas and bloating (7), constipation (14), blood in poop (2), diarrhoea (7), worry about going poop (5), worry about stomach aches (2), medicines (4), communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (acute V.3.0) – parent report for young children (ages 5–7)[54 55] Gastrointestinal symptoms module (acute V.3.0) – young child report (ages 5–7)[54 55]	Total (74): 3-point and 5-point LS. Stomach pain and hurt (6), stomach discomfort when eating (5), food and drink limits (6), trouble swallowing (3), heartburn and reflux (4), nausea and vomiting (4), gas and bloating (7), constipation (14), blood in poop (2), diarrhoea (7), worry about going poop (5), worry about stomach aches (2), medicines (4), communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (acute V.3.0) – parent report for children (ages 8–12)[54 55] Gastrointestinal symptoms module (acute V.3.0) – child report (ages 8–12)[54 55] Gastrointestinal symptoms module (acute V.3.0) – parent report for teens (ages 13–18)[54 55] Gastrointestinal symptoms module (acute V.3.0) – teens report (ages 13–18)[54 55]	Total (74): 3-point and 5-point LS. Stomach pain and hurt (6), stomach discomfort when eating (5), food and drink limits (6), trouble swallowing (3), heartburn and reflux (4), nausea and vomiting (4), gas and bloating (7), constipation (14), blood in poop (bowel movement) (2), diarrhoea (7), worry about going poop (bowel movements) (5), worry about stomach aches (2), medicines (4), communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
(ii) Rome IV			
Rome IV – parent-report form for infants and toddlers (ages 0–3) (R49QG-toddler)[57]	Total (29 for ages 0–12 months; 18 for ages 1–3 years): Infant gastrointestinal problems (11), vomiting (9), bowel movements (9)	Defined diagnostic criteria for functional gastrointestinal disorders in neonates and toddlers:[57] Infant regurgitation, infant rumination syndrome, cyclic vomiting syndrome, infant colic, functional diarrhoea, infant dyschezia, functional constipation.	
Parent-report form for children and adolescents (4 years of age and older) (R4PDQ-child)[58] Self-report form for children and adolescents (10 years of age and older) (R4PDQ-child)[58]	Total (42): Belly ache and uncomfortable feelings above the belly button (12), belly aches and abdominal pain around and below the belly button (10), bowel movements (7), nausea and vomiting (9), other symptoms (4).	Defined diagnostic criteria for functional gastrointestinal disorders in children and adolescents:[58] Cyclic vomiting syndrome, functional nausea and functional vomiting, rumination syndrome, aerophagia, functional dyspepsia, irritable bowel syndrome, abdominal migraine, functional abdominal pain – not otherwise specified, functional constipation, non-retentive faecal incontinence.	
(iii) Spence Children's Anxiety Scale			
Spence – Preschool Anxiety Scale (parent report) (ages 0 to 4)[59 60]	Total (34): 5-point LS. Generalised anxiety (5), social anxiety (6), obsessive compulsive disorder (5), physical injury fears (7), separation anxiety (5).	Responses are scored 0 (not true at all) to 4 (very often true). A maximum possible score of 112.	A score 1 SD above mean for a subscale or total score warrants further clinical investigation. A score of 0.5 SD above the mean on total score is indicative of an elevated, but not clinical level of anxiety.

Continued

Table 2 Continued

Measure	Domains (Items)	Scoring	Interpretation
Spence Children's Anxiety Scale (parent report) (5 years and older)[59 60]	Total (38 scored, 39 total): 4-point LS. Panic attack and agoraphobia (9), separation anxiety (6), physical injury fears (5), social phobia (6), obsessive compulsive (6), generalised anxiety disorder/overanxious disorder (6).	Responses are scored 0 (never) to 3 (always). A maximum possible score of 114. T-score calculation.	A score 1 SD above mean (T-score of ≥60) for a subscale or total score is indicative of subclinical or elevated levels of anxiety warranting further clinical investigation.
Spence Children's Anxiety Scale (8 years and older)[59 60]	Total (38 scores, 45 total): 4-point LS. Separation anxiety (6), social phobia (6), obsessive compulsive (6), panic attack and agoraphobia (9), physical injury fears (5), generalised anxiety (6).		
(iv) Short Mood and Feelings Questionnaire			
Mood and Feelings Questionnaire: Short Version (parent report on Child) (ages 6–18)[61 62 75]	Total (13): 3-point LS. Depressive symptoms (13).	Responses are scored 0 (not true) to 2 (true). A maximum possible score of 26.	Higher scores suggest more severe depressive symptoms. A score of ≥12 may indicate the presence of depression in the respondent.
Mood and Feelings Questionnaire: Short Version (child self-report) (ages 6–18)[61 62 75]	Total (13): 3-point LS. Depressive symptoms (13).		

All questionnaires will be collected from all participants at each time point.
N.B. Disease-specific questionnaires can be added into the Qualtrics data collection form, that is, the Paediatric Sleep Questionnaire: Sleep-Disordered Breathing Subscale,[76] for children with OSA.
HRQOL, health-related quality of life; LS, Likert scale; OSA, obstructive sleep apnoea.

are de-identified using the unique study ID. Data will be stored securely as per ethics review board guidelines.

Handling of abnormal outcomes or distress

The well-being of participants is of utmost importance. Participants and their parents/guardians will be advised to contact any of the study investigators if they have concerns regarding any aspects of their participation. It is possible that thinking about one's health or the health of one's child may elicit emotional distress in some participants. Depending on the nature of the concern or level of distress communicated, a relevant study investigator will contact the participant and/or his or her primary caregiver by telephone or in-person to assess any concerns and arrange appropriate follow-up or referral as soon as possible. Additionally, all Participant Information Sheets will provide the details for several, free, age-appropriate 24 hours telephone-based support services. All individuals will be clearly informed that choosing not to take part in the study, or withdrawing from the study at any stage, will not adversely affect their or their child's healthcare or relationship with hospital staff in any way.

Bias, confounding factors and handling of missing data

The single-centre nature of this study is a limitation due to the restricted recruitment pool available and potential for selection bias; however, SCH is a tertiary referral centre for a diverse group of children across the state of New South Wales, which is the most populous state in Australia. Age and gender are known confounding factors for microbiome analyses and are controlled for with matching. There are rapid changes in the intestinal microbiota during the first 3 years of life, after which it becomes relatively stable.[68–70] Although we aim to match participants as closely as possible, our criteria for acceptable matching is as follows:

- ► For children less than 1 year, age to be matched within 3 months,
- ► For children aged 1 to 3 years, age to be matched within 6 months,
- ► For children aged 4 years and older, age to be matched within 2 years.

Additional confounding factors for microbiome analyses include perinatal factors and ethnicity, for which sensitivity analyses will be performed. Condition specific medications (eg, pancreatic enzyme replacement therapy or antibiotic therapy in CF) are potential confounders and attempts to control for these factors will be made at the analysis stage. Vaccines may influence short-term cytokine production and analyses as a confounder will be considered.[71 72] Missing data will be treated as missing and accounted for using linear mixed models (see statistical methods below).

Study size

In an exploratory research programme of this nature, with multiple conditions of interest, sample size calculations for the primary outcomes are difficult. As an

initial, arbitrary target, three males and three females in each of the following age ranges (0 to 5, >5 to 10, >10 to 18 years) will be recruited to account for age-related and gender-related changes in microbiomes and diet. This calculation assumes that six participants will be required for most statistical tests of interest and an analysis can be performed on the smallest subgroup (eg, six CF vs six HC children aged 0 to 5 years). Therefore 18 participants for each condition and 18 HC (which can be used for comparison against multiple conditions) are an initial target sample size. Initial data from this sample size can then be used for subsequent power-focussed study designs.

Statistical methods

Statistical analyses will be performed in R V.3.4.4. All outcome measures will be analysed cross-sectionally and temporally. Descriptive statistics will be calculated for all outcome parameters for each cohort according to normality of distribution. Given the sample size, categorical variables will be compared using Fisher's exact test. Continuous variables will be analysed according to distribution with a Student's t-test or Wilcoxon signed-rank test for parametric and non-parametric data, respectively. A linear random-effects mixed model or variant of generalised linear-mixed model will be used to evaluate cross-sectional and temporal differences in outcome measures. This technique will allow for control of confounders and treatment of missing data as missing. Correlations between two continuous variables will be performed using Pearson or Spearman correlations according to distribution. Alpha diversity indices will be measured by richness (number of taxa) and Shannon index. Phylogeny-based and taxonomy-based beta diversity will be calculated using UniFrac distances[46] and Bray-Curtis dissimilarities, respectively, and used to generate non-metric multidimensional scaling plots. Permutational multivariate analysis of variance tests (permutations=1000) will be used to test if beta diversity significantly differs between groups and age using the vegan function adonis.[73] A significant difference in abundance of taxa, proteins or metabolites between groups will be assessed using the analysis of composition of microbiomes package V.1.1–3.[74] For all analyses, $p < 0.05$ (two-tailed) is considered significant except in the instance of multiple comparisons, in which case a Benjamini & Hochberg correction will be applied and $q < 0.05$ will be considered significant.

ETHICS AND DISSEMINATION

The EARTH research programme received ethics approval from the Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/18/SCHN/26). Any amendment to the protocol which may impact the conduct of the study will be approved by the ethics committee before implementation.

The results of studies from this research programme will be presented in international conferences and will be published in peer-reviewed journals. Findings may also be presented as: (i) easy-to-read summaries for participants and the community; (ii) educational lectures and seminars for patients, families and the community; (iii) website and social media postings; (iv) newsletter updates for study participants; (v) reports for relevant advocacy groups and funding partners.

EXPECTED OUTCOMES AND SIGNIFICANCE OF THE RESEARCH PROJECT

To our knowledge the EARTH research programme will be the first in children with a chronic gastrointestinal and/or respiratory condition to simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes. By exploring disease mechanisms and environmental interactions (ie, diet) we may in turn develop insights into potential therapeutic strategies. Studies from this programme have the potential for direct translation into clinical care as diet is a highly modifiable factor. This programme also provides a structured approach for performing prospective, longitudinal, controlled, observational studies which can be simultaneously applied to multiple health conditions, and utilised a universal control cohort.

Author affiliations
[1]Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia
[2]Department of Physiotherapy, Sydney Children's Hospital Randwick, Sydney, New South Wales, Australia
[3]Department of Respiratory, Sydney Children's Hospital Randwick, Sydney, New South Wales, Australia
[4]Department of Surgery, Sydney Children's Hospital Randwick & Neuroscience Research Australia (NeuRA), Sydney, New South Wales, Australia
[5]School of Medical Sciences, University of New South Wales, Sydney, New South Wales, Australia
[6]Virology Research Laboratory, Prince of Wales Hospital, Sydney, New South Wales, Australia
[7]Molecular and Integrative Cystic Fibrosis (miCF) Research Centre®, Sydney, New South Wales, Australia
[8]Cincinnati Children's Center for Heart Disease and the Developing Mind, Heart Institute and the Division of Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center & Department of Pediatrics, University of Cincinnati, Cincinnati, Ohio, USA
[9]Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia
[10]Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Respiratory, Sydney Children's Hospital Randwick, Randwick, New South Wales, Australia
[11]Department of Nutrition and Dietetics, Sydney Children's Hospital Randwick, Sydney, New South Wales, Australia

ORCID iD
Michael J Coffey http://orcid.org/0000-0002-5543-9438

REFERENCES

1. Burns KH, Casey PH, Lyle RE, et al. Increasing prevalence of medically complex children in US hospitals. *Pediatrics* 2010;126:638–46.
2. Satterwhite BB. Impact of chronic illness on child and family: an overview based on five surveys with implications for management. *Int J Rehabil Res* 1978;1:7–17.
3. Wijlaars LPMM, Gilbert R, Hardelid P. Chronic conditions in children and young people: learning from administrative data. *Arch Dis Child* 2016;101:881–5.
4. , Peterson J, Garges S, et al, NIH HMP Working Group. The NIH human microbiome project. *Genome Res* 2009;19:2317–23.
5. Pinquart M, Shen Y. Depressive symptoms in children and adolescents with chronic physical illness: an updated meta-analysis. *J Pediatr Psychol* 2011;36:375–84.
6. Conviser JH, Fisher SD, McColley SA. Are children with chronic illnesses requiring dietary therapy at risk for disordered eating or eating disorders? A systematic review. *Int J Eat Disord* 2018;51:187–213.
7. Saxby N, Painter C, Kench A. *Nutrition guidelines for cystic fibrosis in Australia and New Zealand*. Sydney: Thoracic Society of Australia and New Zealand, 2017.
8. Statovci D, Aguilera M, MacSharry J, et al. The impact of Western diet and nutrients on the microbiota and immune response at mucosal interfaces. *Front Immunol* 2017;8:838.
9. Valdes AM, Walter J, Segal E, et al. Role of the gut microbiota in nutrition and health. *BMJ* 2018;361:k2179.
10. Farrell PM, Rosenstein BJ, White TB, et al. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: cystic fibrosis Foundation consensus report. *J Pediatr* 2008;153:S4–14.
11. Massie RJ, Olsen M, Glazner J, Robertson CF, et al. Newborn screening for cystic fibrosis in Victoria: 10 years' experience (1989-1998). *Med J Aust* 2000;172:584–7.
12. Sutherland R, Katz T, Liu V, et al. Dietary intake of energy-dense, nutrient-poor and nutrient-dense food sources in children with cystic fibrosis. *J Cyst Fibros* 2018;17:804–10.
13. Nielsen S, Needham B, Leach ST, et al. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. *Sci Rep* 2016;6:24857.
14. Dhaliwal J, Leach S, Katz T, et al. Intestinal inflammation and impact on growth in children with cystic fibrosis. *J Pediatr Gastroenterol Nutr* 2015;60:521–6.
15. Pang T, Leach ST, Katz T, et al. Elevated fecal M2-pyruvate kinase in children with cystic fibrosis: a clue to the increased risk of intestinal malignancy in adulthood? *J Gastroenterol Hepatol* 2015;30:866–71.
16. Ooi CY, Pang T, Leach ST, et al. Fecal human β-defensin 2 in children with cystic fibrosis: is there a diminished intestinal innate immune response? *Dig Dis Sci* 2015;60:2946–52.
17. Lee JM, Leach ST, Katz T, et al. Update of faecal markers of inflammation in children with cystic fibrosis. *Mediators Inflamm* 2012;2012:1–6.
18. Coffey MJ, Nightingale S, Ooi CY. Predicting a biliary aetiology in paediatric acute pancreatitis. *Arch Dis Child* 2013;98:965–9.
19. Coffey MJ, Nightingale S, Ooi CY. Serum lipase as an early predictor of severity in pediatric acute pancreatitis. *J Pediatr Gastroenterol Nutr* 2013;56:602–8.
20. Kaakoush NO, Pickford R, Jaffe A, et al. Is there a role for stool metabolomics in cystic fibrosis? *Pediatr Int* 2016;58:808–11.
21. Debyser G, Mesuere B, Clement L, et al. Faecal proteomics: a tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *J Cyst Fibros* 2016;15:242–50.
22. Kosorok MR, Zeng L, West SE, et al. Acceleration of lung disease in children with cystic fibrosis after Pseudomonas aeruginosa acquisition. *Pediatr Pulmonol* 2001;32:277–87.
23. Nixon GM, Davey M. Sleep apnoea in the child. *Aust Fam Physician* 2015;44:352–5.
24. Moreno-Indias I, Torres M, Montserrat JM, et al. Intermittent hypoxia alters gut microbiota diversity in a mouse model of sleep apnoea. *Eur Respir J* 2015;45:1055–65.
25. Poroyko VA, Carreras A, Khalyfa A, et al. Chronic sleep disruption alters gut microbiota, induces systemic and adipose tissue inflammation and insulin resistance in mice. *Sci Rep* 2016;6:35405.
26. Xue J, Zhou D, Poulsen O, et al. Intermittent hypoxia and hypercapnia accelerate atherosclerosis, partially via Trimethylamine-Oxide. *Am J Respir Cell Mol Biol* 2017;57:581–8.
27. Tripathi A, Melnik AV, Xue J, et al. Intermittent hypoxia and hypercapnia, a hallmark of obstructive sleep apnea, alters the gut microbiome and metabolome. *mSystems* 2018;3. doi:10.1128/mSystems.00020-18
28. Durgan DJ, Ganesh BP, Cope JL, et al. Role of the gut microbiome in obstructive sleep Apnea-Induced hypertension. *Hypertension* 2016;67:469–74.
29. Barceló A, Esquinas C, Robles J, et al. Gut epithelial barrier markers in patients with obstructive sleep apnea. *Sleep Med* 2016;26:12–15.
30. Wu BG, Sulaiman I, Wang J, et al. Severe obstructive sleep apnea is associated with alterations in the nasal microbiome and an increase in inflammation. *Am J Respir Crit Care Med* 2019;199:99–109.
31. Frykman PK, Nordenskjöld A, Kawaguchi A, et al. Characterization of bacterial and fungal microbiome in children with Hirschsprung disease with and without a history of enterocolitis: a multicenter study. *PLoS One* 2015;10:e0124172.
32. Demehri FR, Frykman PK, Cheng Z, et al. Altered fecal short chain fatty acid composition in children with a history of Hirschsprung-associated enterocolitis. *J Pediatr Surg* 2016;51:81–6.
33. Barr JJ, Auro R, Furlan M, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci U S A* 2013;110:10771–6.
34. Zuo T, Lu X-J, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut* 2019;68:1169–79.
35. Minot S, Sinha R, Chen J, et al. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res* 2011;21:1616–25.
36. Rodriguez-Brito B, Li L, Wegley L, et al. Viral and microbial community dynamics in four aquatic environments. *Isme J* 2010;4:739–51.
37. Kan JM, Cowan CSM, Ooi CY, et al. What can the gut microbiome teach us about the connections between child physical and mental health? A systematic review. *Dev Psychobiol* 2019;61:700–13.
38. Dinan TG, Cryan JF. The microbiome-gut-brain axis in health and disease. *Gastroenterol Clin North Am* 2017;46:77–89.
39. He Y, Wen Q, Yao F, et al. Gut-lung axis: the microbial contributions and clinical implications. *Crit Rev Microbiol* 2017;43:81–95.
40. Callaghan BL, Fields A, Gee DG, et al. Mind and gut: associations between mood and gastrointestinal distress in children exposed to adversity. *Dev Psychopathol* 2020;32:309–28.

41 Barfod KK, Roggenbuck M, Hansen LH, *et al*. The murine lung microbiome in relation to the intestinal and vaginal bacterial communities. *BMC Microbiol* 2013;13:303.

42 Chan A-W, Tetzlaff JM, Altman DG, *et al*. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.

43 Wheeler R. Gillick or Fraser? a plea for consistency over competence in children. *BMJ* 2006;332:807.

44 Whittaker RH. Evolution and measurement of species diversity. *Taxon* 1972;21:213–51.

45 Shannon CE. A mathematical theory of communication. *Bell System Technical Journal* 1948;27:379–423.

46 Lozupone CA, Hamady M, Kelley ST, *et al*. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* 2007;73:1576–85.

47 Bray JR, Curtis JT. An Ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 1957;27:325–49.

48 Hubbell SP. *The unified neutral theory of biodiversity and biogeography*. Princeton, N.J: Princeton University Press, 2001.

49 Conceição-Neto N, Zeller M, Lefrère H, *et al*. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 2015;5:16532.

50 Watson JF, Collins CE, Sibbritt DW, *et al*. Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *Int J Behav Nutr Phys Act* 2009;6:62.

51 Burrows T, Berthon B, Garg ML, *et al*. A comparative validation of a child food frequency questionnaire using red blood cell membrane fatty acids. *Eur J Clin Nutr* 2012;66:825–9.

52 Burrows TL, Warren JM, Colyvas K, *et al*. Validation of overweight children's fruit and vegetable intake using plasma carotenoids. *Obesity* 2009;17:162–8.

53 Collins CE, Burrows TL, Truby H, *et al*. Comparison of energy intake in toddlers assessed by food frequency questionnaire and total energy expenditure measured by the doubly labeled water method. *J Acad Nutr Diet* 2013;113:459–63.

54 Varni JW, Bendo CB, Denham J, *et al*. PedsQL gastrointestinal symptoms module: feasibility, reliability, and validity. *J Pediatr Gastroenterol Nutr* 2014;59:347–55.

55 Varni JW, Kay MT, Limbers CA, *et al*. PedsQL gastrointestinal symptoms module item development: qualitative methods. *J Pediatr Gastroenterol Nutr* 2012;54:664–71.

56 Varni JW, Limbers CA, Neighbors K, *et al*. The PedsQL™ infant scales: feasibility, internal consistency reliability, and validity in healthy and ill infants. *Qual Life Res* 2011;20:45–55.

57 Benninga MA, Faure C, Hyman PE, *et al*. Childhood functional gastrointestinal disorders: Neonate/Toddler. *Gastroenterology* 2016;150:1443–55.

58 Hyams JS, Di Lorenzo C, Saps M, *et al*. Functional disorders: children and adolescents. *Gastroenterology* 2016.

59 Nauta MH, Scholing A, Rapee RM, *et al*. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. *Behav Res Ther* 2004;42:813–39.

60 Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *J Abnorm Psychol* 1997;106:280–97.

61 Angold A, Costello EJ, Messer SC, *et al*. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Methods Psychiatr Res* 1995;5:237–49.

62 Messer SC, Angold A, Costello EJ, *et al*. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: factor composition and structure across development. *Int J Methods Psychiatr Res* 1995;5:251–62.

63 Australian Bureau of statistics. *Socio-Economic indexes for areas (SEIFA) technical paper 2016*. Canberra: ABS, 2016.

64 Edgar RC. Search and clustering orders of magnitude faster than blast. *Bioinformatics* 2010;26:2460–1.

65 Lin J, Kramna L, Autio R, *et al*. Vipie: web pipeline for parallel characterization of viral populations from multiple NGS samples. *BMC Genomics* 2017;18:378.

66 Roux S, Enault F, Hurwitz BL, *et al*. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 2015;3:e985.

67 Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;26:1367–72.

68 Yatsunenko T, Rey FE, Manary MJ, *et al*. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.

69 Derrien M, Alvarez A-S, de Vos WM. The gut microbiota in the first decade of life. *Trends Microbiol* 2019;27:997–1010.

70 Milani C, Duranti S, Bottacini F, *et al*. The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. *Microbiol Mol Biol Rev* 2017;81. doi:10.1128/MMBR.00036-17. [Epub ahead of print: 08 Nov 2017].

71 Simon WL, Salk HM, Ovsyannikova IG, *et al*. Cytokine production associated with smallpox vaccine responses. *Immunotherapy* 2014;6:1097–112.

72 Talaat KR, Halsey NA, Cox AB, *et al*. Rapid changes in serum cytokines and chemokines in response to inactivated influenza vaccination. *Influenza Other Respir Viruses* 2018;12:202–10.

73 Anderson JL, Miles C, Tierney AC. Effect of probiotics on respiratory, gastrointestinal and nutritional outcomes in patients with cystic fibrosis: a systematic review. *J Cyst Fibros* 2017;16:186–97.

74 Mandal S, Van Treuren W, White RA, *et al*. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis* 2015;26:27663.

75 Daviss WB, Birmaher B, Melhem NA, *et al*. Criterion validity of the mood and feelings questionnaire for depressive episodes in clinic and non-clinic subjects. *J Child Psychol Psychiatry* 2006;47:927–34.

76 Chervin RD, Hedger K, Dillon JE, *et al*. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep Med* 2000;1:21–32.